Filtering Grouped Data

A brief introduction

Having clause

The HAVING clause is the final clause of the SELECT statement and a way of filtering data that has already been grouped.

Pseudo code example:

SELECT PeopleNames FROM World

WHERE HeightInFeet>6

HAVING Count(Spouse)=0

Order matters

SELECT Customers.CustState, SUM(Order_Details.QuantityOrdered * Order_Details.QuotedPrice) AS SumOfOrders
FROM(Customers INNER JOIN Orders ON Customers.CustomerID = Orders.CustomerID) INNER JOIN Order_Details ON Orders.OrderNumber = Order_Details.OrderNumber
GROUP BY Customers.CustState
HAVING SUM(Order_Details.QuantityOrdered*Order_Details.QuotedPrice) > 1000000
AND Customers.CustState IN('WA', 'OR', 'CA')

Order matters: Explained

The problematic lines are below:

GROUP BY Customers.CustState
HAVING SUM(Order_Details.QuantityOrdered*Order_Details.QuotedPrice) > 1000000
AND Customers.CustState IN('WA', 'OR', 'CA')

**The data is grouped, evaluated and then filtered. This method requires more work of the database than necessary. It is like asking someone to go upstairs count all the books and then asking them to give you the totals for young adult fiction and sci-fi after they have already come back downstairs. The next slide will show the more efficient ordering.

Order matters: Revised

```
SELECT Customers.CustState, SUM(Order_Details.QuantityOrdered *
Order_Details.QuotedPrice) AS SumOfOrders
FROM (Customers INNER JOIN Orders ON Customers.CustomerID =
Orders.CustomerID) INNER JOIN Order_Details ON Orders.OrderNumber =
Order_Details.OrderNumber
WHERE  Customers.CustState IN('WA', 'OR', 'CA')
GROUP BY Customers.CustState
HAVING SUM(Order_Details.QuantityOrdered*Order_Details.QuotedPrice) >
1000000
```

Having or Where clause?

Take a look at the next two slides.

1) What question do you think the query is answering?
2) Which one do you think is correct and why?

Write down your estimations and run the queries using the School Scheduling db.

Examples using Having clause

Use the School Scheduling db to run the query.

```
SELECT Categories.CategoryDescription,
COUNT(Faculty_Categories.StaffID) AS ProfCount
FROM (Categories
INNER JOIN Faculty_Categories
ON Categories.CategoryID = Faculty_Categories.CategoryID)
INNER JOIN Faculty ON Faculty.StaffID = Faculty_Categories.StaffID
WHERE Faculty.Title='Professor'
GROUP BY Categories.CategoryDescription
HAVING COUNT(Faculty_Categories.StaffID)<3
```

Examples using Where clause

Use the School Scheduling db to run the query.

```
SELECT Categories.CategoryDescription, (SELECT
COUNT(Faculty.StaffID) FROM(Faculty INNER JOIN Faculty_Categories
ON Faculty.StaffID = Faculty_Categories.StaffID) INNER JOIN Categories
AS C2 ON C2.CategoryID = Faculty_Categories.CategoryID WHERE
C2.CategoryID = Categories.CategoryID and Faculty.Title='Professor') AS
ProfCount
FROM Categories WHERE(SELECT COUNT(Faculty.StaffID) FROM
(Faculty INNER JOIN Faculty_Categories ON Faculty.StaffID =
Faculty_Categories.StaffID) INNER JOIN Categories AS C3 ON
C3.CategoryID = Faculty_Categories.CategoryID WHERE C3.CategoryID
= Categories.CategoryID AND Faculty.Title = 'Professor')<3
```

Having or Where clause?

HAVING seems to be the star of this presentation but it actually not the correct way of executing this query.

Did you find it strange that there were no rows with a “zero” value? It may be coincidence but as we see it was not. You may not have noticed the repetition if you copied and pasted the query from the previous slide but the subquery is where we were able to look for the total professors(top pg.488) and then compare that to the number of subjects with fewer than three full professors teaching(bottom subquery pg.488).

Still fuzzy, consult pages 485-489 or watch the accompanying video.

Try It - 1 Entertainment db

Show each agent’s name, the sum of the contract price for the engagements booked, and the agent’s total commission for the agents whose total commission is more than $1,000.

Try It - 2 School Scheduling db

Show me the subject categories that have fewer than three full professors teaching that subject.

Try It - 3

List each staff member and the count of the classes each is scheduled to teach for those staff members who have fewer than three classes.

Try It - 4 Recipes db

For what class of recipes are there two or more recipes?